CROSSWALK TERMINOLOGY PROJECT

A collaborative project by the Managing Re-identification Risk (MRR) Working Group

CONTENTS

ACKNOWLEDGMENTS

The Federal Privacy Council (FPC) Managing Re-identification Risk (MRR) Working Group undertook a project to identify and document terms that are relevant to managing re-identification risk, but may be used differently across professional domains. The definitions presented in this document were collected between June and December 2022 through a series of data calls to the following executive councils:

1. Chief Data Officers Council (CDOC)
2. Evaluation Officer Council (EOC)
3. Interagency Council on Statistical Policy (ICSP)
4. Federal Privacy Council (FPC)

ABSTRACT

All federal agencies are challenged to establish or improve the processes for disclosure avoidance in response to the Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act). As a result of the Evidence Act, an understanding of shared terminology is becoming increasingly more critical across the Federal Government. This document represents a collaborative effort to improve communication among information and data governance professionals.

The Crosswalk Terminology Project identified 24 terms that are relevant to managing re-identification risk, but may be used differently across professional domains. As critical partners in this federal agency-wide effort, the following councils contributed to this effort:

1. Chief Data Officers Council (CDOC)
2. Evaluation Officer Council (EOC)
3. Interagency Council on Statistical Policy (ICSP)
4. Federal Privacy Council (FPC)

BACKGROUND

LEADERSHIP

The FPC MRR Working Group operates with the support and direction of the FPC Executive Committee. The working group is co-chaired by Shannan Catalano of ED's Student Privacy Policy Office and Jamie Huang of EPA's Privacy Office.

MEMBERSHIP

The FPC MRR Working Group is a community of practice for professionals who are working with disclosure avoidance programs that function as part of an agency privacy program. Members of the working group include disclosure review program managers, coordinators, disclosure review board chairs, or other persons fulfilling similar substantive and administrative roles within their agency. The working group is open to all federal agency staff interested in the group's focus.

MISSION AND STRATEGY

Informed by the Evidence Act, the FPC MRR Working Group's mission focuses on collaboration and key processes falling within the working group's scope. Mission and scope can evolve over time based on the guidance and recommendations of the FPC leadership.

Mission

- Improve communication among information and data governance professionals.
- Improve upon processes for managing disclosure avoidance and re-identification risk in response to the Evidence Act.

Strategy

- Collaborate with critical federal partners.
- Identify key processes to which the FPC MRR Working Group can contribute.

CHALLENGES

Establishing successful programs within existing privacy frameworks is complicated by two challenges:

- Coordinating with federal agencies among privacy professionals, technical methodologists, and IT professionals requires understanding of shared language to discuss complex disclosure avoidance concepts.
- Within and across federal agencies, disclosure avoidance programs must account for risks associated with the mosaic effect within a broader privacy framework. The mosaic effect is an especially complex issue because it has the potential to occur within and across federal agencies.

These issues are particularly important for disclosure avoidance programs operating with larger privacy programs in their federal agency. The issues are important even to federal departments and agencies not directly responding to Evidence Act requirements because federal staff must still engage with the broader community and professional domains, which include those who are responding to the requirements of the Evidence Act.

CONCLUDED PROJECT

Having a shared language around disclosure avoidance is becoming increasingly critical across the Federal Government. The FPC MRR Working Group completed this project to improve communication among information and data governance professionals by identifying terms that are not only relevant to managing re-identification risk, but also may be used differently across professional domains. Twenty-four terms have been identified.

DATA CALL FOR TERMS/REFERENCES/SOURCES

Based on a data call to the Evidence Act Councils, the FPC MRR Working Group selected the following terms for this project:

Accountability
Anonymization
Availability
Confidentiality
De-identification
Differential Privacy
Direct Identifier
Disclosure Avoidance
Disclosure Limitation
Disassociability
Formal Privacy
Information in Identifiable Form (IIF)
Indirect Identifier
Integrity
Linked or Linkable Information
Manageability
Mathematical Privacy
Mosaic Effect
Personally Identifiable Information (PII)
Predictability
Re-identification
Re-identification Risk
Sensitive PII
Statistical Disclosure Limitation
Other Terms Used (e.g., U.S. Person PII)

TERMS AND DEFINITIONS

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Accountability	FPC	1. (General context) Property that ensures the actions of an entity may be traced uniquely to the entity. 2. (Systems security context) The security objective that generates the requirement for actions of an entity to be traced uniquely to that entity. This supports non-repudiation, deterrence, fault isolation, intrusion detection and prevention, and after action recovery and legal action. 3. (Privacy context) A Fair Information Practice Principle (FIPP) privacy principle that refers to an organization's requirements to demonstrate their implementation of the FIPPs and applicable privacy requirements.	CNSS 4099 from ISO/IEC 7498-2:1989; NIST SP 800-33 (adapted); OMB A-130 (adapted); Committee on National Security Systems (CNSS) Glossary: *https://www.niap-ccevs.org/Ref/CNSSI_4009.pdf*
Accountability	CDOC	—	—
Accountability	EOC	Assigns responsibility for actions or achieving targeted results.	—
Accountability	ICSP	Generally a deprecated term due to ambiguity and contradictory usage in the technical literature.	NISTIR 8053 § 1.4.1, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*
Anonymization	FPC	Process that removes the association between the identifying dataset and the data subject. The description of "anonymization" as a type of de-identification.	NISTIR 8053 § 1.4.1, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/nist.ir.8053.pdf* NIST SP 800-122, Guide to Protecting the Confidentiality of Personally Identifiable Information (PII): *https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-122.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Anonymization	CDOC	The process of making previously identifiable information de-identified and for which a code or other association for re-identification no longer exists. (In medical terms: data from which the patient cannot be identified by the recipient of the information.)	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf* NIST SP 800-122, Guide to Protecting the Confidentiality of Personally Identifiable Information (PII): *https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-122.pdf*
Anonymization	EOC	Process that removes the association between the identifying dataset and the data subject. Anonymized data are data that have been de-identified and do not include a re-identification code. To remove identifying information so the original source cannot be known.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf* Student Privacy Glossary: *https://studentprivacy.ed.gov/glossary*
Anonymization	ICSP	—	—
Attribute Disclosure	FPC	—	—
Attribute Disclosure	CDOC	—	—
Attribute Disclosure	EOC	—	—
Attribute Disclosure	ICSP	Attribute disclosure occurs when an attribute or characteristic about an identified individual (or entity) is revealed through a data release.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Availability	FPC	The term 'availability' means ensuring timely and reliable access to and use of information. The property that data or information is accessible and usable upon demand by an authorized person.	Federal Information Security Modernization Act of 2014 (FISMA 2014), Public Law No: 113-283 (12/18/2014): https://www.congress.gov/bill/113th-congress/senate-bill/2521
Availability	CDOC	Ensuring timely and reliable access to and use of information.	OMB A-130; 44 U.S.C. § 3552, Managing Information as a Strategic Asset: https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf
Availability	EOC	Timely and reliable access to and use of data and results.	Adapted from E-Government Act of 2002: https://www.congress.gov/107/plaws/publ347/PLAW-107publ347.pdf
Availability	ICSP	Timely, reliable access to information or a service.	NIST SP 800-152, A Profile for U.S. Federal Cryptographic Key Management Systems: https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-152.pdf

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Collection of Information	FPC	A. The obtaining, causing to be obtained, soliciting, or requiring the disclosure to third parties or the public of facts or opinions by or for an agency, regardless of form or format, calling for either: i. Answers to identical questions posed to, or identical reporting or recordkeeping requirements imposed on, ten or more persons, other than agencies, instrumentalities, or employees of the United States. ii. Answers to questions posed to agencies, instrumentalities, or employees of the United States that are to be used for general statistical purposes. B. Shall not include a collection of information described under section 3518(c)(1).	44 U.S.C. § 3502(3), Public Printing and Documents: *https://www.govinfo.gov/content/pkg/USCODE-2021-title44/pdf/USCODE-2021-title44-chap35-subchapI-sec3502.pdf* Note that 44 U.S.C. § 3502(3) is a FISMA-specific definition applicable to federal information systems. Under subsection (b), it excludes information described as a "collection of information" under section 44 U.S.C. § 3518(c)(1).
Collection of Information	CDOC	—	—
Collection of Information	EOC	—	—
Collection of Information	ICSP	—	—
Confidentiality	FPC	The term "confidentiality" means preserving authorized restrictions on access and disclosure, including means for protecting personal privacy and proprietary information.	OMB A-130; FIPS 200, 44 U.S.C. § 3542, Minimum Security Requirements for Federal Information and Information Systems: *https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.200.pdf*
Confidentiality	CDOC	Preserving authorized restrictions on access and disclosure, including means for protecting personal privacy and proprietary information.	OMB A-130; 44 U.S.C. § 3552, Managing Information as a Strategic Asset: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Confidentiality	EOC	Preserving authorized restrictions on data access and disclosure, including means for protecting personal privacy and proprietary information and personally identifiable information (PII).	Adapted from E-Government Act of 2002: *https://www.congress.gov/107/plaws/publ347/PLAW-107publ347.pdf*
Confidentiality	ICSP	Confidentiality is the "property that sensitive information is not disclosed to unauthorized entities."	NIST SP 800-152, A Profile for U.S. Federal Cryptographic Key Management Systems: *https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-152.pdf*
Data Subject	FPC	While the term "data subject" is not used in the Privacy Act, see the definition of "individual."	5 U.S.C. § 552a, (a)(2), Public information; agency rules, opinions, orders, records, and proceedings: *https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title5-section552&num=0&edition=prelim*
Data Subject	CDOC	—	—
Data Subject	EOC	—	—
Data Subject	ICSP	A data subject is any person, establishment, or other entity from or about whom data are collected.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1*
Deductive Disclosure	FPC	—	—
Deductive Disclosure	CDOC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Deductive Disclosure	EOC	The identification of an individual's identity using known characteristics of that individual.	Deductive Disclosure Risk: *https://www.icpsr.umich.edu/web/pages/DSDR/disclosure.html#:~:text=Deductive%20disclosure%20is%20the%20identification,identify%20respondents%20with%20unique%20characteristics*
Deductive Disclosure	ICSP	—	—
De-identification	FPC	General term for any process of removing the association between a set of identifying data and the data subject.	NIST SP 800-53, Revision 5, Security and Privacy Controls for Information Systems and Organizations: *https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53r5.pdf*
De-identification	CDOC	—	—
De-identification	EOC	Process of removing the association between a set of identifying data and the data subject. To adhere to promises of confidentiality made during the informed consent process and to mitigate risks to data providers for providing personally identifiable information (PII) and sensitive data in the data package.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf* Adapted from MCC Guidelines for Transparent, Reproducible, and Ethical Data and Documentation (TREDD): *https://www.mcc.gov/resources/pub-full/guidance-mcc-guidelines-tredd*
De-identification	ICSP	The process by which statistical disclosure limitation methods are applied to reduce (not eliminate) the likelihood of disclosure from the resulting data.	Federal Committee on Statistical Methodology's Data Protection Toolkit, The Challenge of De-Identification: *https://nces.ed.gov/fcsm/dpt/content/1-3-2*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Differential Privacy	FPC	A set of techniques based on a mathematical definition of identity disclosure and information leakage from operations on a dataset. Differential privacy prevents disclosure by adding non-deterministic noise (usually small random values) to the results of mathematical operations before the results are reported. A system for publicly sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*
Differential Privacy	CDOC	—	—
Differential Privacy	EOC	A set of techniques based on a mathematical definition of identity disclosure and information leakage from operations on a dataset. Differential privacy prevents disclosure by adding non-deterministic noise (usually small random values) to the results of mathematical operations before the results are reported. A system for publicly sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Differential Privacy	ICSP	Differential privacy is a rigorous mathematical definition of privacy for statistical analysis and machine learning. In the simplest setting, consider an algorithm that analyzes a dataset and releases statistics about it (such as means and variances, cross-tabulations, or the parameters of a machine learning model). Such an algorithm is said to be differentially private if, by looking at the output one cannot tell whether an individual's data was included in the original dataset or not. In other words, the guarantee of a differentially private algorithm is that its behavior hardly changes when a single individual joins or leaves the dataset; anything the algorithm might output on a database containing some individual's information is almost as likely to have come from a database without that individual's information. Most notably, this guarantee holds for every individual and every dataset. Therefore, regardless of how eccentric any single individual's details are, and regardless of the details of anyone else in the database, the guarantee of differential privacy still holds. This gives a formal guarantee that individual-level information about participants in the database is not leaked. Differential privacy achieves this strong guarantee by carefully injecting random noise into computation of the released statistics, so as to hide the effect of each individual.	What is Differential Privacy: *https://opendp.org/about*
Direct Identifier	FPC	While the term "direct identifier" is not explicitly used in the Privacy Act, the term "record" within the statute is used to refer to information about an individual that is maintained by an agency and "contains his name, or the identifying number, symbol, or other identifying particular assigned to the individual, such as a finger or voice print or a photograph." (Related: A "system of records" within the Privacy Act refers to a group of records "from which information is retrieved by the name of the individual or by some identifying number, symbol, or other identifying particular assigned to the individual.")	5 U.S.C. § 552a, (a)(4), and (5), Public information; agency rules, opinions, orders, records, and proceedings: *https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title5-section552&num=0&edition=prelim*
Direct Identifier	CDOC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Direct Identifier	EOC	Data that directly identifies an individual; examples include personally identifiable information (PII) such as names, social security numbers, and email addresses.	Adapted from NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*
Direct Identifier	ICSP	A direct identifier is a data element with a 1:1 relationship to a data subject that can be used to uniquely identify that individual or entity, such as a Social Security Number, Tax Identification Number, or other unique pieces of information. As such, provision of a direct identifier by definition leads to identity disclosure.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1-3-2*
Disclosure	FPC	While the Privacy Act does not explicitly define what "disclosure" means, it uses the term in the context of sharing "any record which is contained in a system of records by any means of communication to any person, or to another agency."	5 U.S.C. § 552a (b), Public information; agency rules, opinions, orders, records, and proceedings: *https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title5-section552&num=0&edition=prelim*
Disclosure	CDOC	—	—
Disclosure	EOC	—	—
Disclosure	ICSP	Disclosure is a concept that can mean different things in different contexts. Within the context of the Data Protection Toolkit, disclosure entails the release or exposure of information that was supposed to be confidential. Disclosure of confidential information about a data subject can take one of three forms: identity disclosure, attribute disclosure, and inferential disclosure.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1*
Disclosure Avoidance	FPC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Disclosure Avoidance	CDOC	—	—
Disclosure Avoidance	EOC	Efforts made to de-identify data in order to reduce the risk of disclosure of personally identifiable information (PII).	Student Privacy Glossary: *https://studentprivacy.ed.gov/glossary*
Disclosure Avoidance	ICSP	Colloquial term for "Statistical Disclosure Limitation."	—
Disclosure Limitation	FPC	—	—
Disclosure Limitation	CDOC	—	—
Disclosure Limitation	EOC	A technique used to manipulate the data prior to release to minimize the risk of inadvertent or unauthorized disclosure of personally identifiable information (PII).	Student Privacy Glossary: *https://studentprivacy.ed.gov/glossary*
Disclosure Limitation	ICSP	See "Statistical Disclosure Limitation."	—
Disclosure Risk	FPC	—	—
Disclosure Risk	CDOC	—	—
Disclosure Risk	EOC	—	—
Disclosure Risk	ICSP	The likelihood that: 1) an individual's (or entity's) identity can be determined through linkage to an external data source (identity disclosure); and 2) an attribute can be learned or better inferred about the individual (attribute or inferential disclosure).	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1-3-4*
Disassociability	FPC	—	—
Disassociability	CDOC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Disassociability	EOC	Enabling the processing of personally identifiable information (PII) or events without association to individuals or devices beyond the operational requirements of the system.	NISTIR 8062, An Introduction to Privacy Engineering and Risk Management in Federal Systems: *https://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8062.pdf*
Disassociability	ICSP	—	—
Dissemination	FPC	It should also be noted that "dissemination" is treated as a distinct information action in OMB Circular A-130 (e.g., handling personally identifiable information (PII) involves "creation, collection, use, processing, storage, maintenance, dissemination, disclosure, and disposal"). A-130 defines "dissemination" as "the government-initiated distribution of information to a nongovernment entity, including the public." It explicitly excludes "distribution limited to Federal Government employees, intra- or interagency use or sharing of Federal information, and responses to requests for agency records under the Freedom of Information Act or the Privacy Act" from the definitional scope.	OMB A-130, Managing Information as a Strategic Asset: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf*
Dissemination	CDOC	—	—
Dissemination	EOC	—	—
Dissemination	ICSP	—	—
Formal Privacy	FPC	—	—
Formal Privacy	CDOC	—	—
Formal Privacy	EOC	Not used in the evaluation community.	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Formal Privacy	ICSP	Every data release results in a loss of some privacy, if only in a probabilistic sense by allowing the data user to make slightly more accurate inferences about people, households, or businesses (Dwork and Naor, 2010). A formally private approach quantifies the amount of privacy loss in any release and requires the total privacy loss from a data release not exceed a certain budget. This type of method protects against a worst case with respect to what information a data user has external to the agency-provided data. A larger privacy budget may allow the data provider to release a greater quantity or quality of data but also creates more privacy loss. Furthermore, unlike traditional disclosure avoidance methods, formally private methods must work even if the entire methodology is publicly known, including parameters. Ideally, the code would even be posted publicly, as long as the seeds for random number generators were not revealed. This limitation prevents the user from undoing the disclosure avoidance in any one instance but allows a complete picture of how the data were protected in a probabilistic sense. Differential Privacy is one type of Formal Privacy.	Freiman, Michael H., Rodriguez, Rolando A., Reiter, Jerome P., and Lauger, Amy. "Formal Privacy and Synthetic Data for the American Community Survey": *https://nces.ed.gov/FCSM/pdf/D5FreimanRodriguezReiter.pdf#:~:text=A%20formally%20private%20approach%20quantifies%20the%20amount%20of,data%20user%20has%20external%20to%20the%20agency-provided%20data*.
Identity Disclosure	FPC	—	—
Identity Disclosure	CDOC	—	—
Identity Disclosure	EOC	—	—
Identity Disclosure	ICSP	Identity disclosure occurs when it is possible to directly identify a specific individual (or entity) in a data release.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Individual/ U.S. Person	FPC	The term "individual" means a citizen of the United States or an alien lawfully admitted for permanent residence.	5 U.S.C. § 552a, (a)(2), Public information; agency rules, opinions, orders, records, and proceedings: *https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title5-section552&num=0&edition=prelim*
Individual/ U.S. Person	CDOC	—	—
Individual/ U.S. Person	EOC	—	—
Individual/ U.S. Person	ICSP	—	—
Inferential Disclosure	FPC	—	—
Inferential Disclosure	CDOC	—	—
Inferential Disclosure	EOC	—	—
Inferential Disclosure	ICSP	Inferential disclosure occurs when information about a data subject can be inferred with high confidence from statistical properties of the data.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Information in Identifiable Form (IIF)	FPC	Any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either direct or indirect means. Information for data collection for program evaluation purposes: i. That directly identifies an individual (e.g., name, address, social security number or other identifying number or code, telephone number, email address) or ii. By which an agency intends to identify specific individuals in conjunction with other data elements, i.e., indirect identification. (These data elements may include a combination of gender, race, birth date, geographic indicator, and other descriptors.)	E-Government Act of 2002 § 208(d) (Public Law 107-347, 44 U.S.C. § 3501): *https://www2.ed.gov/policy/gen/leg/foia/om-6-108.pdf* OMB M-03-22, OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2003/m03_22.pdf*
Information in Identifiable Form (IIF)	CDOC	—	—
Information in Identifiable Form (IIF)	EOC	Any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either direct or indirect means. Information for data collection for program evaluation purposes: i. That directly identifies an individual (e.g., name, address, social security number or other identifying number or code, telephone number, email address) or ii. By which an agency intends to identify specific individuals in conjunction with other data elements, i.e., indirect identification. (These data elements may include a combination of gender, race, birth date, geographic indicator, and other descriptors.)	E-Government Act of 2002 § 208(d) (Public Law 107-347,44 U.S.C. § 3501): *https://www2.ed.gov/policy/gen/leg/foia/om-6-108.pdf* OMB M-03-22, OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2003/m03_22.pdf*
Information in Identifiable Form (IIF)	ICSP	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Indirect Identifier	FPC	While the term "indirect identifier" is not used in OMB Circular A-130, see definition of "Personally Identifiable Information."	—
Indirect Identifier	CDOC	—	—
Indirect Identifier	EOC	Information that can be used to identify an individual through association (or combination) with other information; examples include place of birth, birth date, and geographic indicators such as zip codes.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf* Student Privacy Glossary: *https://studentprivacy.ed.gov/glossary*
Indirect Identifier	ICSP	An indirect identifier is a data element that may be combined with other variables on the file or other external information to identify a specific data subject. Name and birthdate are good examples of indirect identifiers that can easily be used in this manner. There may be numerous John Smiths in a given population, and there may be many individuals that share the same birthdate in that population, but there may be only one John Smith with a particular birthdate. As such, the provision of indirect identifiers can lead to identify, attribute, and inferential disclosure.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1-3-2*
Integrity	FPC	Guarding against improper information modification or destruction, including ensuring information non-repudiation and authenticity.	OMB A-130, Managing Information as a Strategic Asset: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf*
Integrity	CDOC	Guarding against improper information modification or destruction, including ensuring information non-repudiation and authenticity.	OMB A-130, Managing Information as a Strategic Asset: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/circulars/A130/a130revised.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Integrity	EOC	The protection of evaluation data and results from unauthorized access or revision to ensure the data is not compromised through corruption or falsification. Safeguards against improper data modification or destruction, including ensuring data non-repudiation and authenticity. Guarding against improper data modification or destruction, and including ensuring data non-repudiation and authenticity.	Adapted from: Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies, 67 FR 8452 (Feb. 22, 2002): *https://www.federalregister.gov/documents/2002/02/22/R2-59/guidelines-for-ensuring-and-maximizing-the-quality-objectivity-utility-and-integrity-of-information* The Privacy Act and Personally Identifiable Information (5 FAM 460): *https://fam.state.gov/FAM/05FAM/05FAM0460.html* 44 U.S.C. § 3552: *https://www.govinfo.gov/content/pkg/USCODE-2020-title44/pdf/USCODE-2020-title44-chap35-subchapII-sec3552.pdf* Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices (OMB M-20-12): *https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf* Improving Implementation of the Information Quality Act (OMB M-19-15): *https://www.whitehouse.gov/wp-content/uploads/2019/04/M-19-15.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Integrity	ICSP	A property whereby data has not been altered in an unauthorized manner since it was created, transmitted, or stored.	NIST SP 800-152, A Profile for U.S. Federal Cryptographic Key Management Systems: *https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-152.pdf*
Linked or Linkable Information	FPC	"Linked or linkable information" is incorporated in the definition of "Personally Identifiable Information": Information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual.	OMB A-130, Managing Information as a Strategic Resource: *https://www.whitehouse.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf*
Linked or Linkable Information	CDOC	—	—
Linked or Linkable Information	EOC	Linked information is information about or related to an individual that is logically associated with other information about the individual. Linkable information is information about or related to an individual for which there is a possibility of logical association with other information about the individual.	NIST SP 800-122, Guide to Protecting the Confidentiality of Personally Identifiable Information (PII): *https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf*
Linked or Linkable Information	ICSP	Information that is associated with a particular data subject or that can be associated with a particular data subject by linkage via Direct Identifiers, Indirect Identifiers, or Pseudo-identifiers.	—
Manageability	FPC	Providing the capability for granular administration of data, including alteration, deletion, and selective disclosure.	NIST Privacy Framework (Jan 2020): *https://www.nist.gov/privacy-framework/privacy-framework*
Manageability	CDOC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Manageability	EOC	Providing the capability for granular administration of personally identifiable information (PII), including alteration, deletion, and selective disclosure.	NISTIR 8062, An Introduction to Privacy Engineering and Risk Management in Federal Systems: *https://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8062.pdf*
Manageability	ICSP	—	—
Mathematical Privacy	FPC	—	—
Mathematical Privacy	CDOC	—	—
Mathematical Privacy	EOC	—	—
Mathematical Privacy	ICSP	See "Formal Privacy"	—
Mosaic Effect	FPC	The mosaic effect occurs when the information in an individual dataset, in isolation, may not pose a risk of identifying an individual (or threatening some other important interest such as security) but when combined with other available information, could pose such risk. Before disclosing potential personally identifiable information (PII) or other potentially sensitive information, agencies must consider other publicly available data—in any medium and from any source—to determine whether some combination of existing data and the data intended to be publicly released could allow for the identification of an individual or pose another security concern.	OMB M-13-13, Open Data Policy—Managing Information as an Asset: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2013/m-13-13.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Mosaic Effect	CDOC	The mosaic effect occurs when the information in an individual dataset, in isolation, may not pose a risk of identifying an individual (or threatening some other important interest such as security), but when combined with other available information could pose such risk. Before disclosing potential personally identifiable information (PII) or other potentially sensitive information, agencies must consider other publicly available data—in any medium and from any source—to determine whether some combination of existing data and the data intended to be publicly released could allow for the identification of an individual or pose another security concern.	OMB M-13-13, Open Data Policy—Managing Information as an Asset: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2013/m-13-13.pdf*
Mosaic Effect	EOC	Occurs when the information in an individual dataset used for program evaluation, in isolation, may not pose a risk of identifying an individual (or threatening some other important interest such as security), but when combined with other available information could pose such risk.	Adapted from OMB M-13-13, Open Data Policy—Managing Information as an Asset: *https://project-open-data.cio.gov/policy-memo/*
Mosaic Effect	ICSP	The mosaic effect occurs when the information in an individual dataset, in isolation, may not pose a risk of identifying an individual (or threatening some other important interest such as security), but when combined with other available information could pose such risk. Before disclosing potential personally identifiable information (PII) or other potentially sensitive information, agencies must consider other publicly available data—in any medium and from any source—to determine whether some combination of existing data and the data intended to be publicly released could allow for the identification of an individual or pose another security concern.	OMB M-13-13, Open Data Policy—Managing Information as an Asset: *https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2013/m-13-13.pdf*
Personally Identifiable Information (PII)	FPC	Information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual.	OMB A-130, Managing Information as a Strategic Resource: *https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A130/a130revised.pdf*, also available at *https://www.fpc.gov/resources/glossary/#S*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Personally Identifiable Information (PII)	CDOC	Personally identifiable information (PII) is information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual. Because there are many different types of information that can be used to distinguish or trace an individual's identity, the term PII is necessarily broad. To determine whether information is PII, the agency shall perform an assessment of the specific risk that an individual can be identified using the information with other information that is linked or linkable to the individual. In performing this assessment, it is important to recognize information that is not PII can become PII whenever additional information becomes available—in any medium or from any source—that would make it possible to identify an individual.	OMB A-130, Managing Information as a Strategic Resource: https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A130/a130revised.pdf, also available at *https://www.fpc.gov/resources/glossary/#S*
Personally Identifiable Information (PII)	EOC	Information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual.	OMB A-130, Managing Information as a Strategic Resource: https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A130/a130revised.pdf, also available at *https://www.fpc.gov/resources/glossary/#S*
Personally Identifiable Information (PII)	ICSP	—	—
Privacy Impact Assessment	FPC	An analysis of how information is handled to ensure handling conforms to applicable legal, regulatory, and policy requirements regarding privacy; to determine the risks and effects of creating, collecting, using, processing, storing, maintaining, disseminating, disclosing, and disposing of information in identifiable form in an electronic information system; and to examine and evaluate protections and alternate processes for handling information to mitigate potential privacy concerns. A privacy impact assessment is both an analysis and a formal document detailing the process and the outcome of the analysis.	OMB A-130, 10.a.63: *https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Privacy Impact Assessment	CDOC	—	—
Privacy Impact Assessment	EOC	A decision tool used to identify and mitigate privacy risks associated with the data collected and used in a program evaluation.	—
Privacy Impact Assessment	ICSP	—	—
Pseudo-identifier	FPC	—	—
Pseudo-identifier	CDOC	—	—
Pseudo-identifier	EOC	—	—
Pseudo-identifier	ICSP	A pseudo-identifier is a type of indirect identifier that may not be an obvious candidate for attempting re-identification (e.g., movie ratings), but that can be used in the proper context as a key linking variable to re-identify the individuals in a dataset.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1-3-2*
Pseudonymization	FPC	A particular type of de-identification that both removes the association with a data subject and adds an association between a particular set of characteristics relating to the data subject and one or more pseudonyms. Typically, pseudonymization is implemented by replacing direct identifiers with a pseudonym, such as a randomly generated value. Pseudonymization is the processing of personal information in a manner that renders the personal information no longer attributable to a specific consumer without the use of additional information, provided that the additional information is kept separately and is subject to technical and organizational measures to ensure the personal information is not attributed to an identified or identifiable consumer.	NISTIR 8053 from ISO/TS 25237; ISO 29100 Privacy Framework; ISO 25237 Health Informatics—Pseudonymization, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/nist.ir.8053.pdf*
Pseudonymization	CDOC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Pseudonymization	EOC	—	—
Pseudonymization	ICSP	—	—
Predictability	FPC	Enabling reliable assumptions by individuals, owners, and operators about data and their processing by a system, product, or service.	NIST Privacy Framework (January 2020): *https://www.nist.gov/privacy-framework/privacy-framework*
Predictability	CDOC	—	—
Predictability	EOC	Enabling reliable assumptions by individuals, owners, and operators about data and their processing by a system, product, or service.	NISTIR 8062, An Introduction to Privacy Engineering and Risk Management in Federal Systems: *https://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8062.pdf*
Predictability	ICSP	—	—
Re-identification	FPC	Re-identification is the process of attempting to discern the identities that have been removed from de-identified data. Because an important goal of de-identification is to prevent unauthorized re-identification, such attempts are sometimes called re-identification attacks.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/nist.ir.8053.pdf*
Re-identification	CDOC	—	—
Re-identification	EOC	Any process that re-establishes the relationship between identifying data and an evaluation data subject.	Adapted from NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*
Re-identification	ICSP	Re-identification is the process of attempting to discern the identities that have been removed from de-identified data.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Re-identification Risk	FPC	Though not an explicit definition, the Evidence Act speaks to "risks and restrictions related to the disclosure of personally identifiable information (PII), including the risk that an individual data asset in isolation does not pose a privacy or confidentiality risk but when combined with other available information may pose such a risk."	44 U.S.C. § 3504(b)(6)(A); 44 U.S.C. § 3511(a)(2)(E)(i), Authority and Functions of Director: *https://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title44-section3504&num=0&edition=prelim*
Re-identification Risk	CDOC	—	—
Re-identification Risk	EOC	The risk that de-identified records can be re-identified.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*
Re-identification Risk	ICSP	See "Disclosure Risk"	—
Rule of 10	FPC	—	—
Rule of 10	CDOC	This is the basic rule that Federal Employee Viewpoint Survey (FEVS) uses to mitigate re-identification risk. Basically, in a given data set, the rule prohibits any combination of data fields in a single dataset that, at the row level, has less than 10 individuals in a given grouping when aggregated. So it is not just less than 10 employees, but also if there is a work unit with 20 people and data fields include gender, job series, and race data revealing only 2 Black female statisticians. In this case, one or more columns would be removed to preserve the rule of 10. The other way this is done is to aggregate up until they can't be identified. FEVS (see https://www.doi.gov/pmb/hr/doifevs): "The 2022 FEVS will include: ...Detailed reporting will resume at the lower levels, with offices smaller than 10 employees consolidated with larger offices to preserve anonymity."	—
Rule of 10	EOC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Rule of 10	ICSP	—	—
Secondary Use	FPC	—	—
Secondary Use	CDOC	—	—
Secondary Use	EOC	The use of data, either personally identifiable information (PII) or de-identified, for a research or evaluation use other than the original use for which it was collected.	—
Secondary Use	ICSP	—	—
Sensitive Personally Identifiable Information (PII)	FPC	—	—
Sensitive Personally Identifiable Information (PII)	CDOC	—	—
Sensitive Personally Identifiable Information (PII)	EOC	Personal information that specifically identifies an individual and, if such information is exposed to unauthorized access, may cause harm to that individual at a moderate or high impact level.	The Privacy Act and Personally Identifiable Information (5 FAM 460): *https://fam.state.gov/FAM/05FAM/05FAM0460.html*
Sensitive Personally Identifiable Information (PII)	ICSP	—	—
Statistical Disclosure Limitation	FPC	—	—

TERMINOLOGY	COUNCIL	DEFINITION	REFERENCE SOURCE
Statistical Disclosure Limitation	CDOC	—	—
Statistical Disclosure Limitation	EOC	Statistical Disclosure Limitation is the discipline concerned with the modification of statistical data in order to prevent third parties working with these data from recognizing individuals in the data. Disclosure limitation is listed above.	NISTIR 8053, De-Identification of Personal Information: *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf*
Statistical Disclosure Limitation	ICSP	Agencies rely on a set of techniques that assess the overall disclosure risk of any particular record, variable, or value, then treat the data in one or more ways to reduce the risk of disclosure when making their data assets accessible to data users. Together, these various techniques used to assess and treat the data represent a branch of statistics known as Statistical Disclosure Limitation (SDL) methods. The goal of any application of SDL methods is to reduce the risk of disclosure to an acceptable level, while maximizing the accuracy and fitness-for-use of the resulting data for their intended purposes.	Federal Committee on Statistical Methodology's Data Protection Toolkit: *https://nces.ed.gov/fcsm/dpt/content/1*

(Note: "—" indicates that a definition and reference source were not provided)